

K. Mateck

#24 1645

## RAW SEQUENCE LISTING

DATE: 07/26/2001

PATENT APPLICATION: US/09/189,415A

TIME: 16:43:08

Input Set : A:\4021.app

Output Set: N:\CRF3\07262001\I189415A.raw

ENTERED

```

3 <110> APPLICANT: Finlay, Brett B.
4     Kenny, Brendant
5     Devinney, Rebekah
6     Stein, Marcus
8 <120> TITLE OF INVENTION: HOST RECEPTOR FOR PATHOGENIC BACTERIA
10 <130> FILE REFERENCE: 482112.402
12 <140> CURRENT APPLICATION NUMBER: 09/189,415A
13 <141> CURRENT FILING DATE: 1998-11-10
15 <150> PRIOR APPLICATION NUMBER: 60/065,130
16 <151> PRIOR FILING DATE: 1997-11-12
18 <160> NUMBER OF SEQ ID NOS: 9
20 <170> SOFTWARE: PatentIn Ver. 2.1
22 <210> SEQ ID NO: 1
23 <211> LENGTH: 1920
24 <212> TYPE: DNA
25 <213> ORGANISM: Escherichia coli
27 <400> SEQUENCE: 1
28 cggtgcata ccattacgtc atagtaatat aaaggaacgt gtcaaatttc taaataaaaag 60
29 gatatatgta tgcctattgg taaccttggg aataatgtaa atggcaatca ttttaattccc 120
30 cctgcgcgcg cactaccttc acaaacagac ggcgcggcac ggggaggaac tggatcatcta 180
31 attagctcta caggagcatt aggatctcgt tcattgtttt ctcccctgag aaattctatg 240
32 gctgattctg tcgattccag agatattcca ggacttcccta caaacccatc gaggtctgct 300
33 gcagctacat ctgagacatg cttgtctgga ggatttgaag ttctccatga taaggggccca 360
34 cttgatattc tcaatacgca aattggacct tctgcatttc gtgttgaagt gcaggcagat 420
35 ggtactcatg ccgctattgg agaaaaaaat ggtttggagg ttagcgttac attaatgctc 480
36 caagaatgga gcagcttgca atctattgat actgagggtg aaaacagatt tgtttttacc 540
37 gggggacgtg gcggtagtgg gcatccgatg gtcactgtcg catcagatat cgcggaagct 600
38 cgtacgaaaa tactggccaa attagaccca gacaatcatg gaggacgtca acccaaggac 660
39 gttgatacgc gttctgttgg tgttggcagc gttcgggaa tagatgatgg cgttggttagc 720
40 gaaacccata cttcaacaac aaattccagc gttcgtcag atcctaaatt ctgggtttct 780
41 gtgcgcgcaa ttgctgctgg ttagcgggga ctggcgga caaggtattgc acaggcgttg 840
42 gctttgacac cggaaccgga tgatcctaca accaccgatc ctgatcaggc cgcaaatgct 900
43 gcagaaagtg caacaaaaga tcagttaacg caagaagcat tcaagaacct tgagaaccag 960
44 aaagttaaca tcgatgcgaa cggaaatgct attccgtctg gggaattaaa agatgatatt 1020
45 gttgagcaaa tagcacaaca agctaaagag gctggtgagg tggccagaca gcaggctgtt 1080
46 gaaagcaatg cacaggcgca gcagcgatat gaggatcagc atgccagacg tcaggaggaa 1140
47 ttacagcttt catcggttat tggttacggc ctcagcagtg cattgattgt tgctggggga 1200
48 attggtgctg gtgtaacgac tgcgtccat agacgaaatc agccggcaga acagacaact 1260
49 actacaacaa cacatacggg agtcagcaa cacagcggag ggatacccca gcacaaggtg 1320
50 gcaactgatg cacaaagcgc aagacgcttc tctgatagac gtgattcgca ggggagtgtt 1380
51 gcatcgacac actggtcaga ttccctagc gaagtgggta atccatatgc tgaagtggg 1440
52 ggggctcgga atagtctatc ggctcatcag ccagaagagc atatttatga tgaggctcgt 1500
53 gcagatcctg gttatagcgt tattcagaat ttttcaggga gcggcccagt taccggaagg 1560
54 ttaataggaa ctccagggca aggtatccaa agtacttatg cgcttctggc aaacagcggc 1620
55 ggattgcgtt taggtatggg aggattaacg agtgggtggc agacggcagt aagttctgta 1680
56 aatgccgcac caacgcagg accagtagct ttcgtttaaa tatatctgtg agtatttagt 1740
57 tgagggttggg gtggggtggg ggggcgtttt actagcgtaa atgtttcaga gaacaacgtt 1800

```

## RAW SEQUENCE LISTING

DATE: 07/26/2001

PATENT APPLICATION: US/09/189,415A

TIME: 16:43:08

Input Set : A:\4021.app

Output Set: N:\CRF3\07262001\I189415A.raw

```

58 gcagcatggg taactcttga acttctgtta ttataatcaa ttaagagaaa ttataatgtc 1860
59 atcaagatat gaacttttat tagatagggt tgcggaaaaa attggtgttg gatctatttc 1920
62 <210> SEQ ID NO: 2
63 <211> LENGTH: 549
64 <212> TYPE: PRT
65 <213> ORGANISM: Escherichia coli
67 <220> FEATURE:
68 <221> NAME/KEY: VARIANT
69 <222> LOCATION: (314)
70 <223> OTHER INFORMATION: Xaa = any amino acid
72 <400> SEQUENCE: 2
73 Met Pro Ile Gly Asn Leu Gly Asn Asn Val Asn Gly Asn His Leu Ile
74   1           5           10           15
76 Pro Pro Ala Pro Pro Leu Pro Ser Gln Thr Asp Gly Ala Ala Arg Gly
77           20           25           30
79 Gly Thr Gly His Leu Ile Ser Ser Thr Gly Ala Leu Gly Ser Arg Ser
80   35           40           45
82 Leu Phe Ser Pro Leu Arg Asn Ser Met Ala Asp Ser Val Asp Ser Arg
83   50           55           60
85 Asp Ile Pro Gly Leu Pro Thr Asn Pro Ser Arg Leu Ala Ala Ala Thr
86   65           70           75           80
88 Ser Glu Thr Cys Leu Leu Gly Gly Phe Glu Val Leu His Asp Lys Gly
89           85           90           95
91 Pro Leu Asp Ile Leu Asn Thr Gln Ile Gly Pro Ser Ala Phe Arg Val
92   100           105           110
94 Glu Val Gln Ala Asp Gly Thr His Ala Ala Ile Gly Glu Lys Asn Gly
95   115           120           125
97 Leu Glu Val Ser Val Thr Leu Ser Pro Gln Glu Trp Ser Ser Leu Gln
98   130           135           140
100 Ser Ile Asp Thr Glu Gly Lys Asn Arg Phe Val Phe Thr Gly Gly Arg
101 145           150           155           160
103 Gly Gly Ser Gly His Pro Met Val Thr Val Ala Ser Asp Ile Ala Glu
104           165           170           175
106 Ala Arg Thr Arg Ile Leu Ala Lys Leu Asp Pro Asp Asn His Gly Gly
107           180           185           190
109 Arg Gln Pro Lys Asp Val Asp Thr Arg Ser Val Gly Val Gly Ser Ala
110           195           200           205
112 Ser Gly Ile Asp Asp Gly Val Val Ser Glu Thr His Thr Ser Thr Thr
113           210           215           220
115 Asn Ser Ser Val Arg Ser Asp Pro Lys Phe Trp Val Ser Val Gly Ala
116 225           230           235           240
118 Ile Ala Ala Gly Leu Ala Gly Leu Ala Ala Thr Gly Ile Ala Gln Ala
119           245           250           255
121 Leu Ala Leu Thr Pro Glu Pro Asp Asp Pro Thr Thr Thr Asp Pro Asp
122           260           265           270
124 Gln Ala Ala Asn Ala Ala Glu Ser Ala Thr Lys Asp Gln Leu Thr Gln
125           275           280           285
127 Glu Ala Phe Lys Asn Pro Glu Asn Gln Lys Val Asn Ile Asp Ala Asn
128           290           295           300

```

## RAW SEQUENCE LISTING

DATE: 07/26/2001

PATENT APPLICATION: US/09/189,415A

TIME: 16:43:08

Input Set : A:\4021.app

Output Set: N:\CRF3\07262001\I189415A.raw

130 Gly Asn Ala Ile Pro Ser Gly Glu Leu Xaa Asp Asp Ile Val Glu Gln  
 131 305 310 315 320  
 133 Ile Ala Gln Gln Ala Lys Glu Ala Gly Glu Val Ala Arg Gln Gln Ala  
 134 325 330 335  
 136 Val Glu Ser Asn Ala Gln Ala Gln Gln Arg Tyr Glu Asp Gln His Ala  
 137 340 345 350  
 139 Arg Arg Gln Glu Glu Leu Gln Leu Ser Ser Gly Ile Gly Tyr Gly Leu  
 140 355 360 365  
 142 Ser Ser Ala Leu Ile Val Ala Gly Gly Ile Gly Ala Gly Val Thr Thr  
 143 370 375 380  
 145 Ala Leu His Arg Arg Asn Gln Pro Ala Glu Gln Thr Thr Thr Thr  
 146 385 390 395 400  
 148 Thr His Thr Val Val Gln Gln Gln Thr Gly Gly Ile Pro Gln His Lys  
 149 405 410 415  
 151 Val Ala Leu Met Pro Gln Glu Arg Arg Arg Phe Ser Asp Arg Arg Asp  
 152 420 425 430  
 154 Ser Gln Gly Ser Val Ala Ser Thr His Trp Ser Asp Ser Ser Ser Glu  
 155 435 440 445  
 157 Val Val Asn Pro Tyr Ala Glu Val Gly Gly Ala Arg Asn Ser Leu Ser  
 158 450 455 460  
 160 Ala His Gln Pro Glu Glu His Ile Tyr Asp Glu Val Ala Ala Asp Pro  
 161 465 470 475 480  
 163 Gly Tyr Ser Val Ile Gln Asn Phe Ser Gly Ser Gly Pro Val Thr Gly  
 164 485 490 495  
 166 Arg Leu Ile Gly Thr Pro Gly Gln Gly Ile Gln Ser Thr Tyr Ala Leu  
 167 500 505 510  
 169 Leu Ala Asn Ser Gly Gly Leu Arg Leu Gly Met Gly Gly Leu Thr Ser  
 170 515 520 525  
 172 Gly Gly Glu Thr Ala Val Ser Ser Val Asn Ala Ala Pro Thr Pro Gly  
 173 530 535 540  
 175 Pro Val Arg Phe Val  
 176 545  
 179 <210> SEQ ID NO: 3  
 180 <211> LENGTH: 1723  
 181 <212> TYPE: DNA  
 182 <213> ORGANISM: Escherichia coli  
 184 <400> SEQUENCE: 3  
 185 atgcctattg gtaaccttgg tcataatccc aatgtgaata attcaattcc tctgcacct 60  
 186 ccattacctt cacaacccga cgggtgcagg gggcgtgggc agctcattaa ctctacgggg 120  
 187 cggttgggat ctggtgcgct atttacgcct gtaaggaatt ctatggctga ttctggcgac 180  
 188 aatcgtgccg gtgatgttcc tggacttccg gtaaatccga tgcgcctggc ggcgtctgag 240  
 189 ataacactga atgatggatt tgaagttctt catgatcatg gtccgctcga tactcttaac 300  
 190 aggcagattg gctcttcggt atttcgagtt gaaactcagg aagatggtaa acatattgct 360  
 191 gtcggtcaga ggaatggtgt tgagacctct gttgttttaa gtgatcaaga gtacgctcgc 420  
 192 ttgcagtcca ttgatcctga aggtaaagac aaatttgtat ttactggagg ccgtgggtgt 480  
 193 gctgggcatg ctatgggtcac cgttgcttca gatatcacgg aagcccgcga aaggatactg 540  
 194 gagctgttag agcccaaagg gaccggggag tccaaagggt ctggggagtc aaaaggcggt 600  
 195 ggggagttga gggagtcaaa tagcgggtgcg gaaaacacca cagaaactca gacctcaacc 660  
 196 tcaacttcca gccttcgttc agatcctaaa ctttggttgg cgttggggac tgttgctaca 720

## RAW SEQUENCE LISTING

DATE: 07/26/2001

PATENT APPLICATION: US/09/189,415A

TIME: 16:43:08

Input Set : A:\4021.app

Output Set: N:\CRF3\07262001\I189415A.raw

```

197 ggtctgatag ggttggcggc gacgggtatt gtacaggcgc ttgcattgac gccggagccg 780
198 gatagcccaa ccacgaccga ccctgatgca gctgcaagtg caactgaaac tgcgacaaga 840
199 gatcagttaa cgaaagaagc gttccagaac ccagataatc aaaaagttaa tatcgatgag 900
200 ctcggaatg cgattccgtc aggggtattg aaagatgatg ttgttgcgaa tatagaagag 960
201 caggctaaag cagcaggcga agaggccaaa cagcaagcca ttgaaaataa tgctcaggcg 1020
202 caaaaaaat atgatgaaca acaagctaaa cgccaggagg agctgaaagt ttcacggggg 1080
203 gctggctacg gtcttagtgg cgcattgatt cttggtgggg gaattggtgt tgccgtcacc 1140
204 gctgcgcttc atcgaaaaaa tcagccggta gaacaaacaa caacaactac tactacaact 1200
205 acaactacaa gcgcacgtac ggtagagaat aagcctgcaa ataatacacc tgcacagggc 1260
206 aatgtagata cccctgggtc agaagatacc atggagagca gacgtagctc gatggctagc 1320
207 acctcgtcga ctttctttga cacttcagc atagggaccg tgcagaatcc gtatgctgat 1380
208 gttaaaacat cgctgcatga ttgcagggtg ccgacttcta attctaatac gtctgttcag 1440
209 aatatgggga atacagattc tgttgtatat agcaccattc aacatcctcc ccgggatact 1500
210 actgataacg gcgcacgggtt attaggaat ccaagtgcgg ggattcaaag cacttatgcy 1560
211 cgtctggcgc taagtgggtg attacgccat gacatgggag gattaacggg ggggagtaat 1620
212 agcgtgtgta atacttcgaa taaccacca gcgcgggat cccatcgttt cgtctaaata 1680
213 tatcataat cattttattt agaggagggg aggggggaag tct 1723

```

216 &lt;210&gt; SEQ ID NO: 4

217 &lt;211&gt; LENGTH: 559

218 &lt;212&gt; TYPE: PRT

219 &lt;213&gt; ORGANISM: Escherichia coli

221 &lt;400&gt; SEQUENCE: 4

```

222 Met Pro Ile Gly Asn Leu Gly His Asn Pro Asn Val Asn Asn Ser Ile
223   1           5           10           15
225 Pro Pro Ala Pro Pro Leu Pro Ser Gln Thr Asp Gly Ala Gly Gly Arg
226           20           25           30
228 Gly Gln Leu Ile Asn Ser Thr Gly Pro Leu Gly Ser Arg Ala Leu Phe
229           35           40           45
231 Thr Pro Val Arg Asn Ser Met Ala Asp Ser Gly Asp Asn Arg Ala Ser
232           50           55           60
234 Asp Val Pro Gly Leu Pro Val Asn Pro Met Arg Leu Ala Ala Ser Glu
235           65           70           75           80
237 Ile Thr Leu Asn Asp Gly Phe Glu Val Leu His Asp His Gly Pro Leu
238           85           90           95
240 Asp Thr Leu Asn Arg Gln Ile Gly Ser Ser Val Phe Arg Val Glu Thr
241           100          105          110
243 Gln Glu Asp Gly Lys His Ile Ala Val Gly Gln Arg Asn Gly Val Glu
244           115          120          125
246 Thr Ser Val Val Leu Ser Asp Gln Glu Tyr Ala Arg Leu Gln Ser Ile
247           130          135          140
249 Asp Pro Glu Gly Lys Asp Lys Phe Val Phe Thr Gly Gly Arg Gly Gly
250          145          150          155          160
252 Ala Gly His Ala Met Val Thr Val Ala Ser Asp Ile Thr Glu Ala Arg
253           165          170          175
255 Gln Arg Ile Leu Glu Leu Leu Glu Pro Lys Gly Thr Gly Glu Ser Lys
256           180          185          190
258 Gly Ala Gly Glu Ser Lys Gly Val Gly Glu Leu Arg Glu Ser Asn Ser
259           195          200          205
261 Gly Ala Glu Asn Thr Thr Glu Thr Gln Thr Ser Thr Ser Thr Ser

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/189,415A

DATE: 07/26/2001

TIME: 16:43:08

Input Set : A:\4021.app

Output Set: N:\CRF3\07262001\I189415A.raw

```

262      210      215      220
264 Leu Arg Ser Asp Pro Lys Leu Trp Leu Ala Leu Gly Thr Val Ala Thr
265 225      230      235      240
267 Gly Leu Ile Gly Leu Ala Ala Thr Gly Ile Val Gln Ala Leu Ala Leu
268      245      250      255
270 Thr Pro Glu Pro Asp Ser Pro Thr Thr Thr Asp Pro Asp Ala Ala Ala
271      260      265      270
273 Ser Ala Thr Glu Thr Ala Thr Arg Asp Gln Leu Thr Lys Glu Ala Phe
274      275      280      285
276 Gln Asn Pro Asp Asn Gln Lys Val Asn Ile Asp Glu Leu Gly Asn Ala
277      290      295      300
279 Ile Pro Ser Gly Val Leu Lys Asp Asp Val Val Ala Asn Ile Glu Glu
280 305      310      315      320
282 Gln Ala Lys Ala Ala Gly Glu Glu Ala Lys Gln Gln Ala Ile Glu Asn
283      325      330      335
285 Asn Ala Gln Ala Gln Lys Lys Tyr Asp Glu Gln Gln Ala Lys Arg Gln
286      340      345      350
288 Glu Glu Leu Lys Val Ser Ser Gly Ala Gly Tyr Gly Leu Ser Gly Ala
289      355      360      365
291 Leu Ile Leu Gly Gly Gly Ile Gly Val Ala Val Thr Ala Ala Leu His
292      370      375      380
294 Arg Lys Asn Gln Pro Val Glu Gln Thr Thr Thr Thr Thr Thr Thr
295 385      390      395      400
297 Thr Thr Thr Ser Ala Arg Thr Val Glu Asn Lys Pro Ala Asn Asn Thr
298      405      410      415
300 Pro Ala Gln Gly Asn Val Asp Thr Pro Gly Ser Glu Asp Thr Met Glu
301      420      425      430
303 Ser Arg Arg Ser Ser Met Ala Ser Thr Ser Ser Thr Phe Phe Asp Thr
304      435      440      445
306 Ser Ser Ile Gly Gly Pro Cys Arg Ile Arg Met Leu Met Leu Lys His
307      450      455      460
309 Arg Cys Met Ile Arg Arg Cys Arg Leu Leu Ile Leu Ile Arg Leu Phe
310 465      470      475      480
312 Arg Ile Trp Gly Ile Gln Ile Ser Val Val Tyr Ser Thr Ile Gln His
313      485      490      495
315 Pro Pro Arg Asp Thr Thr Asp Asn Gly Ala Arg Leu Leu Gly Asn Pro
316      500      505      510
318 Ser Ala Gly Ile Gln Ser Thr Tyr Ala Arg Leu Ala Leu Ser Gly Gly
319      515      520      525
321 Leu Arg His Asp Met Gly Gly Leu Thr Gly Gly Ser Asn Ser Ala Val
322      530      535      540
324 Asn Thr Ser Asn Asn Pro Pro Ala Pro Gly Ser His Arg Phe Val
325 545      550      555
328 <210> SEQ ID NO: 5
329 <211> LENGTH: 1460
330 <212> TYPE: DNA
331 <213> ORGANISM: Escherichia coli
333 <400> SEQUENCE: 5
334 aattctgttg ctgatgctgc tgattctcgt gccagtgata ttcccgact tcctacaaat 60

```

VERIFICATION SUMMARY

DATE: 07/26/2001

PATENT APPLICATION: US/09/189,415A

TIME: 16:43:09

Input Set : A:\4021.app

Output Set: N:\CRF3\07262001\I189415A.raw

L:130 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2